ENVIRONMENTAL RESEARCH
LETTERS

LETTER • OPEN ACCESS

The realized warming fraction: a multi-model sensitivity study

Environmental Research Letters

LETTER

The realized warming fraction: a multi-model sensitivity study

Patrik L Pfister[1,2] and Thomas F Stocker[1,2]

[1] Climate and Environmental Physics, Physics Institute, University of Bern, 3012 Bern, Switzerland
[2] Oeschger Center for Climate Change Research, University of Bern, 3012 Bern, Switzerland

E-mail: pfister@climate.unibe.ch

Keywords: climate models, ensemble simulations, realized warming fraction, equilibrium climate sensitivity, physical equilibration, energy balance, ocean heat uptake

Supplementary material for this article is available online

OPEN ACCESS

RECEIVED
18 April 2018

REVISED
25 October 2018

ACCEPTED FOR PUBLICATION
26 October 2018

PUBLISHED
17 December 2018

Abstract

The degree of physical-biogeochemical equilibration of the climate system determines for how long global warming will continue after anthropogenic CO_2 emissions have ceased. The physical part of this equilibration process is quantified by the realized warming fraction (RWF), but RWF estimates differ strongly between different climate models. Here we analyze the RWF spread and its physical causes in three model ensembles: 1. an ensemble of comprehensive climate models, 2. an ensemble of reduced-complexity models, and 3. an observationally constrained parameter ensemble of the Bern3D-LPX reduced-complexity model. We show that RWF is generally lower in models with higher equilibrium climate sensitivity. The RWF uncertainty from applying different extrapolation methods for climate sensitivity is substantial, but smaller than the inter-model spread in the three ensembles. We decompose the inter-model spread of RWF using a diagnostic global energy balance model, to compare the spread contribution by the climate sensitivity to contributions by other physical quantities: the efficiency and efficacy of ocean heat uptake, and the effective radiative forcing. In the ensembles of the comprehensive climate models and the Bern3D-LPX model, the spread of the RWF is mostly determined by the spread of the climate sensitivity; for the reduced-complexity models, the spread contribution by the ocean heat uptake efficiency is dominant. Compared to the comprehensive models, the reduced-complexity models have a lower range of climate sensitivities and lower, more unitary ocean heat uptake efficacies, resulting in higher RWF. However, by tuning such models to higher climate sensitivities, they can also achieve RWF values in the lower range of comprehensive models, as demonstrated for Bern3D-LPX. This suggests that reduced-complexity models remain useful tools for future climate change projections, but should employ a range of climate sensitivity tunings to account for the uncertainty in both the long-term warming and the RWF.

1. Introduction

Transient global warming due to greenhouse gas radiative forcing is substantially reduced by ocean heat uptake. However, the fraction of equilibrium warming that is realized in transient climate model simulations differs strongly between models (Winton *et al* 2010, Frölicher and Paynter 2015, Ehlert and Zickfeld 2017).

The realized warming fraction (RWF) (Stouffer 2004, Solomon *et al* 2009) is an important policy-relevant quantity, because models with a lower RWF indicate that global warming may continue for centuries after greenhouse gas emissions cease (Solomon *et al* 2009, Matthews and Zickfeld 2012, Frölicher *et al* 2014, Frölicher and Paynter 2015, Ehlert and Zickfeld 2017). This continued warming is commonly referred to as zero emission warming commitment (ZEC) and may strongly influence long-term climate change mitigation policies. Frölicher and Paynter (2015) show a strong anticorrelation between the RWF and the ZEC in Earth system models (ESMs) from the Climate Model Intercomparison Project phase 5 (CMIP5), and a weaker but consistent anticorrelation in ESMs of Intermediate Complexity (EMICs). This indicates that the ZEC is generally higher in models with a low RWF. Although the ZEC is also influenced by the state of

biogeochemical equilibration, Ehlert and Zickfeld (2017) show that the influence of the physical equilibration (measured by the RWF) is dominant.

The purpose of this study is to investigate the sources of RWF spread among different climate models. Based on the above considerations and references, this is an important step towards understanding the ZEC spread. However, it is more straightforward than investigating the ZEC spread itself, because RWF purely measures the physical model response; carbon cycle uncertainties can be disregarded. Furthermore, RWF can be diagnosed from shorter transient simulations that are available from most models, i.e. idealized future projections where CO_2 concentration increases at a rate of 1% per year.

RWF is generally scenario- and time-dependent (Ehlert and Zickfeld 2017), but here we focus on the RWF at the time of CO_2 doubling. The transient and equilibrium warming at CO_2 doubling are commonly referred to as transient climate response (TCR) and equilibrium climate sensitivity (ECS), respectively (IPCC 2013). Therefore, the RWF at CO_2 doubling is equal to TCR/ECS. TCR and ECS estimates are available for most models, but their comparability is complicated by the heterogeneity of ECS estimation methods (Pfister and Stocker 2017, this study). Earlier studies on CMIP3 and CMIP5 found that the TCR and ECS spread is dominated by the spread in the total climate feedback, with notable secondary influences from radiative forcing and ocean heat uptake for TCR (Dufresne and Bony 2008, Geoffroy *et al* 2012). We investigate how important these influences become for the ratio TCR/ECS, i.e. RWF.

A focus of our investigation is the relation between RWF and the ECS. It has been demonstrated analytically (Hansen *et al* 1984), in a box model (Siegenthaler and Oeschger 1984) and in CMIP3 models (Raper *et al* 2002) that a higher ECS generally causes a lower transient RWF. While widely accepted, a systematic analysis of this finding in more recent models is missing to our knowledge.

We analyze comprehensive ESMs from the CMIP5 ensemble (Taylor *et al* 2012) as well as ESMs of Intermediate Complexity (EMICs) from the EMIC-AR5 ensemble (Eby *et al* 2013). Frölicher and Paynter (2015) have argued that EMICs may be less suitable than ESMs to simulate long-term warming, because they generally have a higher RWF. Such a generalization is questionable considering the heterogeneity of the EMIC ensemble (Ehlert and Zickfeld 2017). To shed more light on this issue, we investigate the causes of the RWF discrepancy between ESMs and EMICs. Furthermore, we also analyze a constrained parameter ensemble of the Bern3D-LPX model (Steinacher *et al* 2013, Steinacher and Joos 2016) and perform four new illustrative simulations with this model.

This paper is structured as follows. Section 2 describes the energy balance framework to understand RWF spread contributions in three published model ensembles and the illustrative Bern3D-LPX simulations. Section 3 presents the relation between the RWF and ECS, first by example of the Bern3D-LPX model and then in the different model ensembles. In section 4, the ensemble spreads of other energy balance parameters, and their relative contributions to the RWF spread are analyzed. We conclude with section 5.

2. Methods

2.1. Energy balance framework

We quantify sources of RWF spread using a global energy balance model (EBM) (Winton *et al* 2010):

$$-\lambda \Delta T = R - \varepsilon N, \tag{1}$$

where ΔT is the warming with respect to preindustrial temperature, R is the radiative forcing, N is the ocean heat uptake and λ is the climate feedback parameter. λ accounts for physical and biogeochemical processes that act to amplify or dampen an existing temperature perturbation, e.g. retreating sea-ice or changing cloud patterns. The global mean EBM is a vast simplification of a time- and location-dependent system, but it enables meaningful intercomparison of more complex models and first-order future projections based on observational data. The remaining EBM parameter ε is the ocean heat uptake efficacy (Winton *et al* 2010), which measures the relative temperature response to a unit N compared to a unit R. $\varepsilon > 1$ implies that the cooling caused by an ocean heat uptake of 1 Wm^{-1} is stronger than the warming caused by a CO_2 forcing of 1 Wm^{-1}. This is the case in many climate models, because the ocean heat uptake and its changes predominantly take place in the high latitudes where local feedbacks are strongest (Winton *et al* 2010, 2013, Armour *et al* 2013, Rose *et al* 2014). Note that the net effect of the forcing is still a warming, because the excess ocean heat uptake in response to the forcing is always smaller than the forcing itself.

Assuming a constant λ, $\varepsilon(t)$ is generally time-dependent, accounting for the non-linearity between ΔT and N (e.g. Paynter and Frölicher 2015). However, calculating $\varepsilon(t)$ from equation (1) may include time-dependencies in $\varepsilon(t)$ that are not due to the ocean heat uptake efficacy (Armour *et al* 2013, Pfister 2017), such as actual time-dependencies in λ (Gregory *et al* 2015, Rose and Rayborn 2016). A more general version of the EBM would thus include both $\varepsilon(t)$ and $\lambda(t)$, but it is difficult to partition the full time-dependence between the two parameters (Pfister 2017, section S1.3 available online at stacks.iop.org/ERL/13/124024/mmedia).

We analyze the RWF in this energy balance framework (equation (1)) by substituting N for $\gamma \Delta T$, where γ is the ocean heat uptake efficiency (Raper *et al* 2002, Kuhlbrodt and Gregory 2012):

$$\mathrm{RWF} = \frac{1}{1 - \varepsilon\gamma/\lambda}, \tag{2a}$$

$$\mathrm{RWF} = \frac{R}{R + \varepsilon\gamma\ \mathrm{ECS}}. \quad (2b)$$

In this view, the RWF for a given R is lowered by higher ε, γ and ECS (or a less negative λ). If we fix these three quantities instead, a higher R increases the RWF.

2.2. Estimation of contributions to the spread of the RWF

The RWF spread across models of a given ensemble is characterized by the inter-model variance σ^2. To obtain an estimate for σ^2 and its uncertainty from the multi-model data, we use a Bayesian calculation (Oliphant 2006). This yields a variance estimator $v = \frac{1}{n-3}\sum_{i=1}^{n}(\mathrm{RWF}_i - \overline{\mathrm{RWF}})^2$ for each model ensemble, where n is the number of ensemble members, RWF_i the RWF of member i and $\overline{\mathrm{RWF}}$ the ensemble mean. For the constrained Bern3D-LPX ensemble, each member RWF_i is weighted by its normalized skill score in reproducing observations (Steinacher *et al* 2013), in both $\overline{\mathrm{RWF}}$ and v calculations. For the EMIC and ESM ensemble, members are not weighted.

How do inter-model differences in the EBM parameters ε, γ, ECS and R contribute to the RWF spread? We compute the Bayesian estimator v_p for the ensemble variance of parameter p analogously to v. Given v_p, we estimate the contribution of this parameter's spread to the RWF spread as

$$v(p) = \left(\frac{\partial\ \mathrm{RWF}}{\partial p}\right)^2 v_p. \quad (3)$$

We apply two different spread decompositions. The first is based on equation (2*b*) and the four parameters ECS, ε, γ and R:

$$\begin{aligned} v_1 &= v(\mathrm{ECS}) + v(\varepsilon) + v(\gamma) + v(R) \\ &= \left(\frac{\mathrm{RWF}^2}{R}\right)^2 \Big((\varepsilon\gamma)^2\ v_{\mathrm{ECS}} + (\gamma\ \mathrm{ECS})^2\ v_\varepsilon \\ &\quad + (\varepsilon\ \mathrm{ECS})^2\ v_\gamma + \left(\frac{\varepsilon\gamma \mathrm{ECS}}{R}\right)^2 v_R \Big). \end{aligned} \quad (4)$$

The second is based on equation (2*a*) and the two parameters λ and $\varepsilon\gamma$. The product $\varepsilon\gamma$ is here treated as a single parameter, which accounts for the total energy balance impact of ocean heat uptake.

$$v_2 = v(\lambda) + v(\varepsilon\gamma) = \left(\frac{\mathrm{RWF}^2}{\lambda}\right)^2 ((\varepsilon\gamma/\lambda)^2\ v_\lambda + v_{\varepsilon\gamma}), \quad (5)$$

where parameters appear as scaling factors for the variances in equations (4) and (5) (e.g. $\left(\frac{\mathrm{RWF}^2}{\lambda}\right)^2$ for $v_{\varepsilon\gamma}$), ensemble mean values of these factors are used (e.g. $\overline{\left(\frac{\mathrm{RWF}^2}{\lambda}\right)^2} \cdot v_{\varepsilon\gamma}$).

The sum of all parameter contributions (i.e. v_1 and v_2) may differ from v if the parameters are not statistically independent. However, this is also true if the contributions are estimated using a more comprehensive analysis of variance (ANOVA, Geoffroy *et al* 2012). We investigate the independence assumption using two different methods. Firstly, we look at parameter cross-correlations and test their statistical significance (table S2). Secondly, the Bayesian variance estimation also yields an uncertainty range for v (Oliphant 2006). If this overlaps with v_1 and v_2, the sum of parameter contributions is in statistical agreement with the total variance (section S1.2).

While the parameters γ, R and ECS are each diagnosed independently, ε is calculated from equation (2*b*) and is thus dependent on these parameters by definition. Its contribution to the RWF spread can therefore be regarded as the variance contribution by the interplay of these parameters causing a non-unitary ocean heat uptake efficacy. As ε is a function of four quantities (including RWF), the correlation of ε with each other independent parameter (γ, R and ECS) can nevertheless be insignificant (table S2).

To test our variance decomposition method, we have also decomposed TCR variance analogously to RWF variance. This can be directly compared to the ANOVA results from Geoffroy *et al* (2012), which are in satisfactory agreement as discussed in section S1.1.

2.3. Analysis of published ESM ensembles

We analyze three different published model ensembles. The first two are a nine-member subset of EMICs from the EMIC-AR5 ensemble (Eby *et al* 2013, section S2.1) and a 15-member subset of comprehensive ESMs from the CMIP5 ensemble (Andrews *et al* 2015, Gregory *et al* 2015, section S2.2). We simply refer to these subsets as 'EMICs' and 'ESMs' in the following. The third ensemble is a constrained parameter ensemble (Steinacher *et al* 2013, Steinacher and Joos 2016, section S2.3) of the Bern3D-LPX model (section 2.4).

For the ESMs, Gregory *et al* (2015) diagnosed TCR and γ as 20 year averages around the time of CO_2 doubling, in an experiment where CO_2 concentration increases by 1% yr^{-1} and stabilizes at 4 × CO_2. We applied the same transient diagnosis for the EMICs and the Bern3D-LPX ensemble, but with a 10 year averaging window for Bern3D-LPX where the available experiment already stabilizes at 2 × CO_2 concentration (section S2.3.2). For all ensembles, we calculated ε from equation (2*b*).

ECS was estimated using the Gregory method (Gregory *et al* 2004) consistently in all three ensembles. Simulation years 150–1000 were extrapolated for the EMICs (Pfister and Stocker 2017) and the Bern3D-LPX ensemble (this study), while the shorter ESM simulations only allow an extrapolation of simulation years 20–150 (Andrews *et al* 2015). While 2 × CO_2 experiments were used for the EMICs and Bern3D-LPX, only 4 × CO_2 experiments are available from CMIP5, which may bias the ECS estimates for state-dependent ESMs (Good *et al* 2015, Pfister and Stocker 2017). This difference in forcing magnitude

may also affect the estimates of the effective radiative forcing R.

R estimates were obtained otherwise consistently for ESMs and EMICs, by a linear fit of the radiative imbalance over simulation years 1–20 back to zero warming (Andrews *et al* 2015, Pfister and Stocker 2017). This estimation is not possible for the Bern3D-LPX ensemble, due to lack of abrupt-forcing simulations. R estimates for this ensemble were therefore obtained by scaling a mean R with the prescribed prior CO_2 forcing scaling (Steinacher *et al* 2013, section S2.3.3). The mean R was obtained from the Bern3D-LPX simulation of EMIC-AR5, consistent with the other EMICs. The R spread of the Bern3D-LPX ensemble may thus be underestimated, as the influence of parameter variations other than prescribed forcing scaling is disregarded.

The above-summarized analyses and differences between ensembles are described in more detail in the supplementary material.

2.4. Illustrative Bern3D-LPX simulations

The Bern3D-LPX model is an EMIC, consisting of a frictional geostrophic ocean and a one-layer moist EBM (Ritz *et al* 2011), coupled to a dynamic global vegetation model (Stocker *et al* 2013). We use an updated model version (Roth *et al* 2014), which mainly has a higher poleward resolution than the version that was used both in EMIC-AR5 and the constrained Bern3D-LPX ensemble by Steinacher *et al* (2013).

The ECS of this model is tuned to 3.0 °C using a global feedback parameter λ_{B3D}. A term $-\lambda_{B3D}\Delta T$, where ΔT is the global mean temperature anomaly from preindustrial, is added to the global energy balance to account for feedbacks that are not otherwise parameterized. This includes cloud feedbacks, but we note that a non-global feedback parameterization would be required to properly emulate the radiative impact of clouds (Ullman and Schmittner 2017). By retuning λ_{B3D}, we create four Bern3D-LPX versions with an ECS of 2.0, 3.0, 4.5 and 6.0 °C, respectively. Simulations with these model versions serve to illustrate the isolated impact of ECS or global mean feedback changes on the RWF.

3. The influence of ECS on the RWF

3.1. Illustrative simulations

Figure 1 shows the simulated temperature evolution of the Bern3D-LPX model in response to an idealized CO_2 scenario following Frölicher and Paynter (2015). CO_2 concentration is first prescribed to increase at a rate of 1% per year until simulation year 99. Thereafter, CO_2 emissions are set to zero and the concentration is allowed to evolve freely, decreasing in response to carbon uptake by the terrestrial and oceanic reservoirs (not shown). Four Bern3D-LPX versions tuned to different ECS are presented (section 2.4).

The temperature evolutions for different ECS deviate not only in magnitude, but also in shape. For low ECS, temperature peaks shortly after the cessation of CO_2 emissions, and the ZEC is small. For high ECS, warming peaks roughly 300 years later and the ZEC is substantial (roughly 1.7 °C for the highest ECS). This implies that both the ZEC and the time until peak warming increase with ECS. Both findings point to an ECS-dependency of the RWF, as the RWF dominantly determines the ZEC (Ehlert and Zickfeld 2017). This ECS-dependency is investigated in the following subsections.

3.2. Ensemble differences and ECS uncertainty

Figure 2 compares these Bern3D-LPX simulations (colors) to corresponding simulations from selected models from the EMIC-AR5 (gray) and CMIP5 (black) ensembles. In contrast to our main analysis presented below, figure 2 uses model data from Frölicher and Paynter (2015) to investigate the RWF-discrepancy that they have pointed out in their selection of EMICs and ESMs.

In agreement with earlier studies (Hansen *et al* 1984, Raper *et al* 2002), figure 2 shows that the RWF is generally lower in models with higher ECS. As ECS is substantially higher in some ESMs than in all of the EMICs, this physical relation partly explains the finding of Frölicher and Paynter (2015) that the RWF is lower in the ESMs compared to the EMICs as a group. This RWF difference is therefore partly related to the tuneable ECS model parameter, and may not be inherent to model complexity. This is further explored in section 4.

Figure 2 also presents the influence of the ECS estimation uncertainty on the RWF. Because equilibrium simulations are not available for most models, ECS has to be estimated using extrapolation methods in abrupt-forcing simulations. For the ESMs, these estimates were obtained using the Gregory method (Gregory *et al* 2004), linearly extrapolating simulation years 20–150 to equilibrium (Andrews *et al* 2015). For the EMICs, 1000 year temperatures are used as ECS estimates (Eby *et al* 2013).

Applying the same estimation methods to our illustrative Bern3D-LPX simulations (colored circles and diamonds in figure 2) underestimates the true model ECS (squares corresponding to 5000 year equilibrium warming). In the two lower-ECS versions of Bern3D-LPX, the values based on year 20–150 Gregory extrapolation (circles) underestimate the true ECS slightly more than the 1000 year temperatures (diamonds), because the global feedback changes still substantially between years 150 and 1000 (Pfister 2017). The opposite is true for other EMICs where feedback changes are small (Pfister and Stocker 2017).

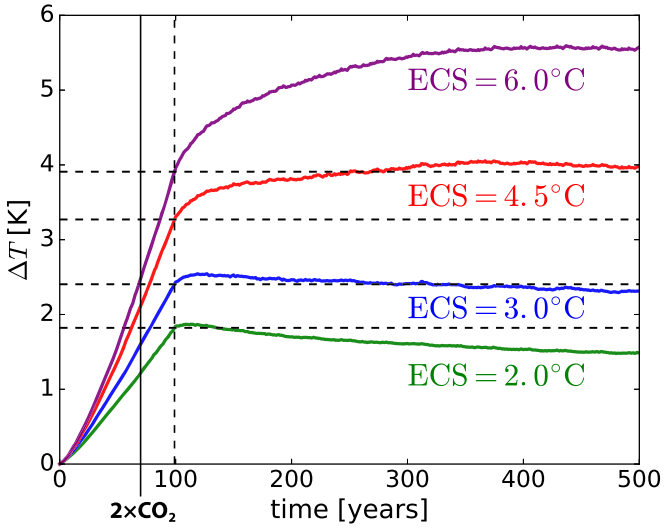

Figure 1. Time series of temperature anomalies in four Bern3D-LPX versions with ECS of 1.5, 3.0, 4.5 and 6.0 K. Scenario: 1% yr^{-1}-CO_2 concentration increase up to year 99 (dashed vertical line), zero CO_2 emissions thereafter (following Frölicher and Paynter 2015). The time of CO_2 doubling (year 70) is marked by a solid vertical line, the warming at year 99 by a dashed horizontal line for each scenario.

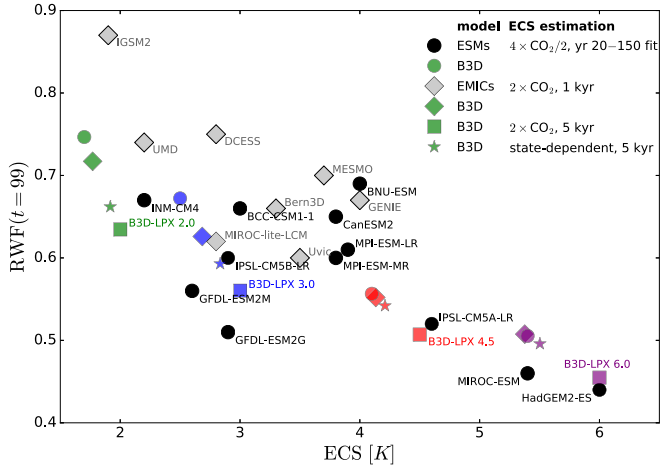

Figure 2. Anticorrelation between ECS and RWF, in year 99 of a 1%-CO_2 scenario in the Bern3D-LPX model (colors as in figure 1) and in different model ensembles as analyzed by Frölicher and Paynter (2015) (EMICs in gray, ESMs in black). For Bern3D-LPX, different ECS estimates are shown: circles and diamonds are diagnosed consistently with ESMs and EMICs, respectively; squares and stars are more accurate equilibrium estimates (see legend and text).

The differences in RWF estimates arising from different ECS estimation methods are smaller than, but of comparable magnitude to, RWF differences between different models. This highlights the fact that RWF estimates can only be as accurate as their underlying ECS estimates.

Furthermore, if the RWF is estimated for any other point in time other than CO_2 doubling, the state-dependence of ECS (e.g. Jonko *et al* 2013, Good *et al* 2015, Pfister and Stocker 2017) has to be taken into account (section S1.3). For the Bern3D-LPX model, equilibrium warming per unit forcing decreases with increasing forcing (Pfister and Stocker 2017). Therefore, the model's actual equilibrium warming corresponding to the forcing at year 99 (stars in figure 2) is smaller than the forcing-scaled ECS (squares), resulting in a higher RWF estimate.

The above considerations complicate the inter-ensemble RWF comparison of Frölicher and Paynter (2015). To avoid these complications, our following RWF analysis focuses on the time of CO_2 doubling, and ECS estimates are consistently obtained in all ensembles by linear extrapolation of the longest available time-series (section 2.3). Only for the illustrative Bern3D-LPX simulations, equilibrium ECS values are available.

3.3. Multi-ensemble analysis and the influence of ocean heat uptake

Figure 3 shows a similar analysis as figure 2, but for a different model selection that is used for the main analysis in this study (methods). Also, the RWF is evaluated in the year of CO_2 doubling (TCR/ECS) as

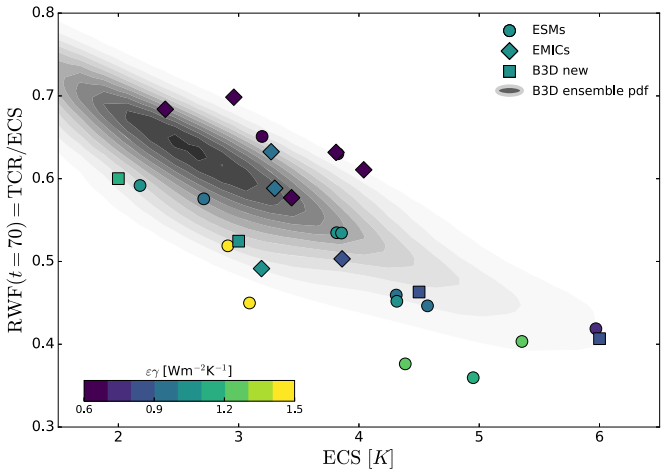

Figure 3. Like figure 2, but RWF is shown at the time of CO_2 doubling (year 70) and ECS estimates are obtained consistently (section 2.3). The same Bern3D-LPX simulations (squares) and different selections of EMICs (diamonds) and ESMs (circles) are shown, see text. The filling color indicates the $\varepsilon\gamma$ value of each model according to the color scale. In addition, the two-dimensional probability density of a constrained Bern3D-LPX ensemble (Steinacher *et al* 2013) is shown in gray shading.

motivated above. Under these slightly different choices of models and time, the finding that some ESMs have a higher ECS and a lower RWF than all EMICs still holds. This is also true when comparing the full CMIP5 and EMIC-AR5 ensembles (table S1). However, the median ECS and RWF differences are smaller than in both model selections presented in figures 2 and 3. Most notably, the median ECS is about 0.4 °C lower in CMIP5 than in both presented ESM selections, indicating that those selections are biased towards ESMs with high ECS. In contrast, our EMIC selection does not bias the median of the RWF and other parameters, but narrows their spreads compared to the full EMIC-AR5 ensemble (table S1).

The filling color of all model markers in figure 3 indicates the product $\varepsilon\gamma$ (equation (2*b*)), that is the influence of transient ocean heat uptake on the RWF. For models with a similar ECS, the RWF is lower for higher $\varepsilon\gamma$ (e.g. the ESMs MIROC5 and NorESM1-M filled in yellow) and vice versa (CanESM2, CNRM-CM5 and the majority of EMICs filled in purple). However, the RWF is also influenced by the forcing R, which explains why $\varepsilon\gamma$ differs substantially between some models with similar ECS and RWF values. The relative importance of ECS, ε, γ and R is examined in section 4.

In addition, figure 3 also shows the two-dimensional probability density function (pdf) of the Bern3D-LPX ensemble. Darker shades of gray imply a larger density of ensemble members, i.e. a larger probability that the constrained model simulates these ECS and RWF values. The shape of the shading reveals a strong ECS/RWF anticorrelation ($r = -0.87$). It is consistent with the other model ensembles ($r = -0.70$ for ESMs and $r = -0.46$ for EMICs, table S2), but the probability density is very low in the region of high-ECS, low-RWF ESMs.

4. The influence of other energy balance parameters

4.1. Ensemble uncertainties of the energy balance parameters

We now compare the influence of ECS on the RWF to the influence of other EBM parameters (equations (2*a*), (2*b*)). We first present inter-model and inter-ensemble differences in those parameters (figure 4) and then analyze their relative contributions to the RWF spread (figure 5).

The most striking parameter difference between EMICs and ESMs apart from the previously investigated RWF and ECS differences (figures 4(a), (d)) is the markedly lower ocean heat uptake efficacy ε of EMICs (figure 4(b)). According to equation (2*b*), this lower ε contributes to the higher RWF of EMICs along with the lower ECS. While the ECS discrepancy to ESMs could be amended by simple model tuning as demonstrated in the new Bern3D-LPX simulations, the difference in ε may be more fundamental and related to model complexity.

The non-unitarity of ε in ESMs has mostly been attributed to changes in the patterns and climate feedbacks of clouds (Rose *et al* 2014, Andrews *et al* 2015, Rose and Rayborn 2016) as well as a related reduction of tropospheric stability (Ceppi and Gregory 2017). While changing temperature patterns acting on constant polar-amplified feedback patterns can also cause non-unitary ε in specific models (Armour *et al* 2013, Pfister 2017), the time-dependency of the feedback pattern generally needs to be taken into account (Rose *et al* 2014).

Only two of the nine EMICs included in our analysis feature a quasi three-dimensional atmosphere and interactive cloud parameterizations, namely both CLIMBER versions. The other EMICs cannot simulate the aforementioned processes responsible for the non-unitarity of ε in ESMs. However, also those

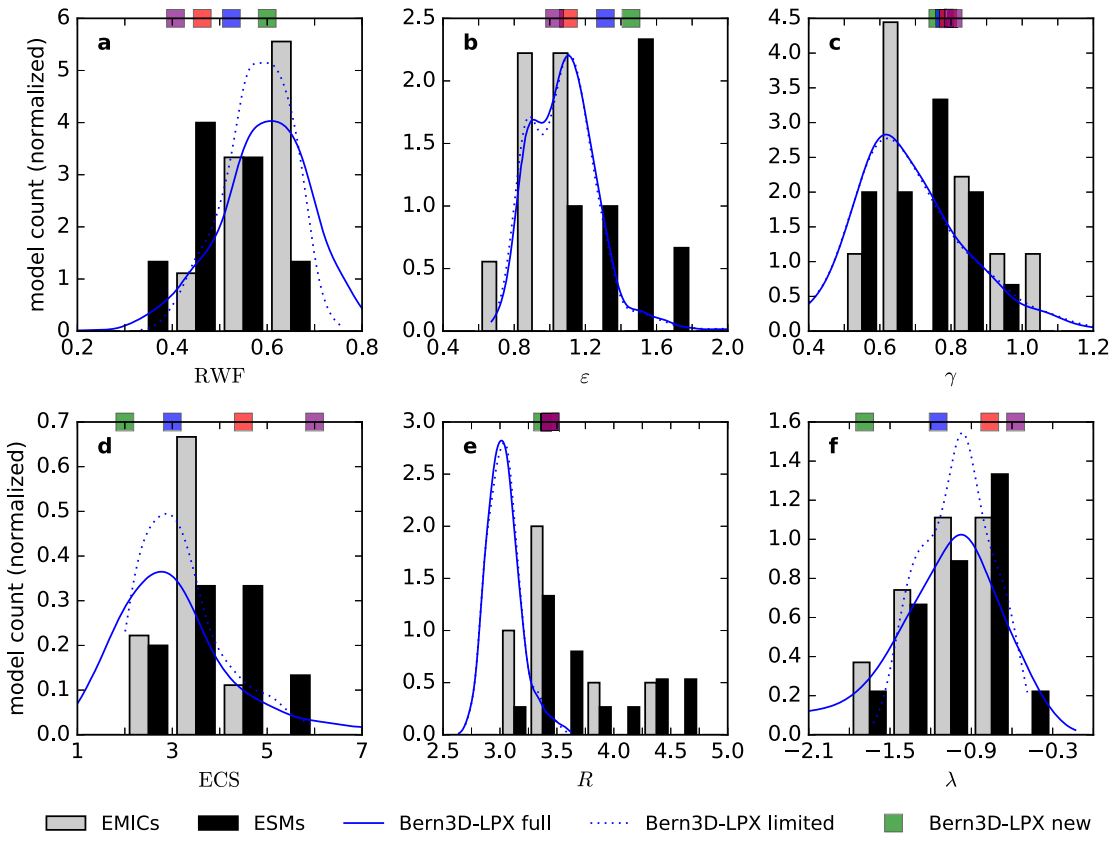

Figure 4. EBM parameters diagnosed from different model ensembles. EMICs and ESMs are shown as side-by-side histograms (gray and black), with bin widths given by x axis ticks. The weighted probability density functions (pdfs) of the constrained Bern3D-LPX ensemble are shown in blue; solid for the full ensemble and dotted for a reduced ECS range of 2 °C–6 °C. All pdfs (and histograms) are normalized such that the area under the pdf (and the sum of bar heights multiplied by widths) amounts to one. Separate parameter values of the four new Bern3D-LPX simulations are shown in colors corresponding to figure 1.

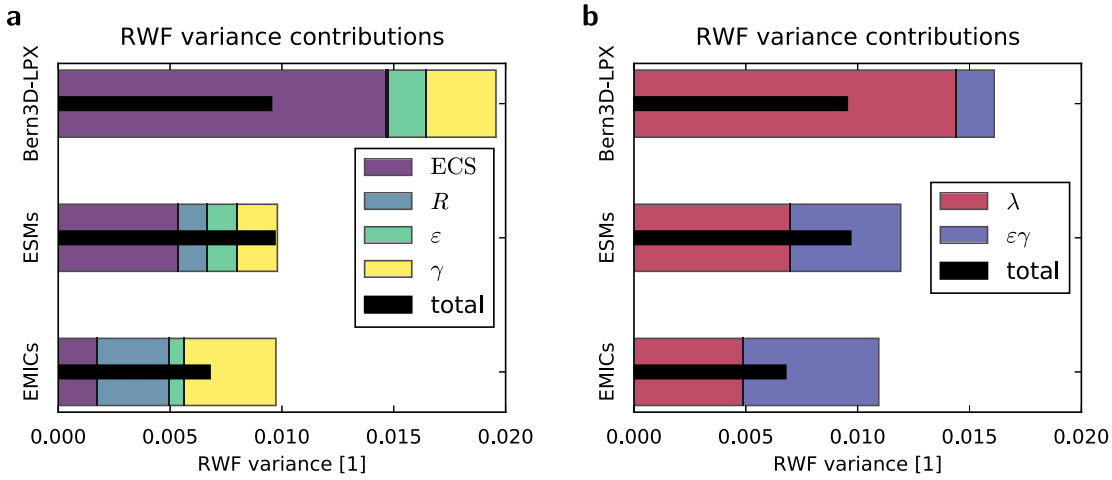

Figure 5. Parameter decomposition of the RWF spread. The total spread (black) is the variance of the RWF in both panels. Colored bars indicate contributions to the RWF spread by the isolated spread of separate EBM parameters, as indicated by the legend; e.g. ν (ECS) for ECS (equation (3)). These are calculated using (a) equation (4) and (b) equation (5).

EMIC-AR5 models that feature interactive cloud parameterizations have an efficacy close to one (table S3). This includes the CLIMBER models and three of the models that have been excluded from our analysis due to restart offset corrections (section S2.1). The inability of these models to simulate non-unitary ε could be explained by coarse atmospheric resolution or other model differences to ESMs.

The only two EMICs simulating substantially non-unitary ε are LOVECLIM ($\varepsilon = 1.45$), which is not

included in the main EMIC analysis due to offset corrections (table S3), and the new version of Bern3D-LPX used for our four illustrative ECS simulations ($\varepsilon = 1.32$ in the model version with ECS = 3.0 °C). In this Bern3D-LPX version, $\varepsilon > 1$ is due to shifting temperature patterns acting on a strong sea-ice albedo feedback (Pfister 2017). LOVECLIM also features a strong sea-ice albedo feedback and no interactive clouds (Eby *et al* 2013), suggesting that a similar effect may be responsible for its non-unitary efficacy. This sea-ice amplification effect is stronger in Bern3D-LPX than in most ESMs (Pfister 2017).

The most likely parameter values inferred from the constrained Bern3D-LPX ensemble—given by the maximum of the pdf—are roughly consistent with the most likely values inferred from the EMICs histogram. An exception to this is R, which is lower for the Bern3D-LPX ensemble. This difference may fully or partly be due to the different diagnosis method of R (section 2.3). Furthermore, some Bern3D-LPX members reach ε values substantially larger than one, which are not found in the EMICs. We hypothesize that these high efficacies may be due to a stronger or more delayed Southern Ocean warming in these members, which was found to cause $\varepsilon > 1$ in the new Bern3D-LPX version (Pfister 2017).

Also for the other EBM parameters, the spreads are generally wider for the Bern3D-LPX ensemble than for the EMICs. Reducing the ECS range to 2 °C–6 °C (dotted pdfs in figure 4) also reduces the spreads of the RWF and of λ, which then become more similar to the corresponding spreads of the EMICs. The distributions of the other EBM parameters are almost unaffected, which indicates that these parameters are reasonably independent of ECS, as confirmed by low correlations (table S2).

4.2. Parameter contributions to the spread of the RWF

How does the spread in each of the EBM parameters (ECS, R, ε, γ, λ) affect the spread of the RWF? In the four illustrative new Bern3D-LPX simulations, it is evident from figure 4 (colored squares) that the RWF is uniquely driven by ECS: the RWF is lower under higher ECS, even though ε is lower (due to faster sea-ice melting, Pfister 2017). Independently, the lower ε should increase the RWF, but this effect is overruled by the stronger ECS effect. γ and R are only very slightly affected by the ECS tuning.

For the three analyzed model ensembles, the answer to this question is less straightforward. We investigate this by calculating separate spread contributions for each parameter, as described in section 2.2. The results are shown in figure 5.

The sum of spread contributions is within the uncertainty of the total spread for the EMICs and ESMS, but not for the constrained Bern3D-LPX ensemble (section S1.2). This is consistent with the fact that there are no significant cross-correlations between parameters (ECS, ε and γ) in the EMICs and ESMs, but significant cross-correlations with γ in Bern3D-LPX (table S2).

The spread discrepancy for Bern3D-LPX is partly resolved by taking the spread of the product $\varepsilon\gamma$ (figure 5(b)), as this is smaller than the sum of separate spreads in ε and γ (figure 5(a)) due to an anticorrelation between those parameters (table S2). The remaining discrepancy may be due to a weak but significant anticorrelation between $\varepsilon\gamma$ and λ (section S1.2). We also note that the discrepancy originates from the long tails of the ECS distribution, as it is resolved if the ensemble is reduced to an ECS range of 2 °C–6 °C (not shown). In the following, we refer to the sum of contributions as 'full RWF spread' to make quantitative statements about relative contributions.

The ECS spread explains most of the RWF spread in the Bern3D-LPX ensemble: it amounts to 75% of the full RWF spread (figure 5(a)), or even to 89% if the dependency of ε and γ is accounted for by taking their product. The ECS spread explains about 55% of the full spread in the ESM ensemble and only 18% in the EMIC ensemble. This shows that the inter-model differences in the RWF mainly originate from the large ECS spread in the Bern3D-LPX and ESM ensembles, but not in the EMIC ensemble that has a smaller ECS spread.

The other EBM parameters contribute in roughly equal parts to the RWF spread of ESMs. In the EMICs, the contribution of ε is smallest, because most models have a near-unitary ocean heat uptake efficacy. For the constrained Bern3D-LPX ensemble, the R spread is very small (Methods) and the ε and γ spreads should not be interpreted separately due to their substantial cross-correlation (table S2).

In the second decomposition (figure 5(b)), we compare the relative spread contributions of the total feedback λ ($= -R/\mathrm{ECS}$) and $\varepsilon\gamma$. In the constrained Bern3D-LPX ensemble, the R spread is small, and thus the relative spread contribution of λ closely corresponds to the spread contribution of ECS (89% of the full spread). This dominant contribution indicates that the degree of equilibration in this ensemble is largely determined by the global feedback tuning. In the ESMs and EMICs, λ and the combined effect of ε and γ are similarly influential for the RWF spread; λ is more influential for the ESMs (58%) and less so for the EMICs (44%), due to the larger ECS (and λ) spread in the ESMs.

5. Conclusions

We have analyzed three ensembles of recent climate models: a subset of 15 ESMs from CMIP5 (Taylor *et al* 2012), a subset of 9 EMICs from EMIC-AR5 (Eby *et al* 2013), and a large observationally constrained parameter ensemble from the Bern3D-LPX model (Steinacher *et al* 2013). All ensembles show that the RWF is lower for

models with higher equilibrium climate sensitivities (ECS). This is a confirmation of analytical considerations (Hansen *et al* 1984) and results from earlier model generations (Raper *et al* 2002, Winton *et al* 2010). We have reproduced this influence of the ECS on the RWF in new Bern3D-LPX simulations using a simple global feedback tuning parameter.

The RWF uncertainty from using different ECS extrapolation methods is substantial in our Bern3D-LPX simulations, but smaller than the intermodel RWF spread in the EMIC and ESM ensembles. For transient RWF estimation at CO_2 concentrations differing from $2 \times CO_2$, the state-dependence of the ECS (e.g. Jonko *et al* 2013, Pfister and Stocker 2017) needs to be taken into account, as demonstrated mainly in the Supplementary section S1.3 of this study.

We have decomposed the RWF spread of each ensemble into contributions from four diagnostic energy balance parameters: ECS, effective radiative forcing, ocean heat uptake efficiency and ocean heat uptake efficacy. In the ESMs and the constrained Bern3D-LPX ensemble, the influence of the ECS spread on the RWF is dominant, explaining 55% and 89% of the RWF spread, respectively. For the ESMs, the remaining parameters contribute about evenly to the RWF spread (13%–19% each). In the EMICs, the smaller ECS spread explains only 18% of the RWF spread, while the dominant contributor is the ocean heat uptake efficiency (42%).

Finally, we have investigated why the RWF tends to be higher in EMICs than in ESMs (Frölicher and Paynter 2015). In our model selection, we have identified lower ECS and ocean heat uptake efficacy ranges in the EMICs compared to the ESMs, which can both contribute to their lower RWF. The lower ECS range of EMICs is confirmed in the model selection of Frölicher and Paynter (2015) and the full EMIC-AR5 and CMIP5 ensembles. However, the median difference of both the RWF and the ECS is smaller in the full ensembles, mainly because the ESM subsets analyzed in our study (following Gregory *et al* 2015) and in Frölicher and Paynter (2015) are biased towards high-ECS ESMs.

We agree with Frölicher and Paynter (2015) that models with low RWF are required to project an upper limit of long-term future warming. However, our study shows that this does not, in principle, rule out the use of EMICs: the RWF spread of the ESMs is mainly caused by their ECS spread, and the ECS of EMICs can be tuned to achieve similarly low RWF values as simulated by ESMs.

As demonstrated for the Bern3D-LPX model, EMIC model versions with a lower RWF can be constructed simply by tuning ECS using a global feedback parameter. Therefore, we suggest that future EMIC studies should sample a range of ECS tunings to account for the uncertainty not only in long-term warming, but also in the degree of physical equilibration measured by the RWF. Massive probabilistic parameter ensembles such as the Bern3D-LPX ensemble by Steinacher *et al* (2013) include such an ECS sampling in addition to other parameters, and therefore remain useful tools for future projections. However, the relative RWF spread contributions of energy balance parameters may differ between such probabilistic ensembles and multi-model ensembles: in the probabilistic Bern3D-LPX ensemble, the ECS spread affects RWF more dominantly than in other ensembles.

For studies where not just the global mean equilibration, but also the patterns of feedbacks and warming are of interest, the discrepancy between EMICs and ESMs is not as easily resolved. The lower ocean heat uptake efficacy range of EMICs compared to ESMs is probably due to the lack of cloud feedbacks in most EMICs. Therefore, it should be investigated whether a cloud feedback emulator (Ullman and Schmittner 2017) would be able to increase the ocean heat uptake efficacy of EMICs, and further contribute to a lower RWF.

Acknowledgments

We thank two anonymous reviewers, and T Frölicher, for their constructive comments that have improved our paper. We acknowledge financial support by the Swiss National Science Foundation through projects 200020_159563 and 200020_172745.

ORCID iDs

Patrik L Pfister https://orcid.org/0000-0002-0390-4844
Thomas F Stocker https://orcid.org/0000-0003-1245-2728

References

Andrews T, Gregory J M and Webb M J 2015 The dependence of radiative forcing and feedback on evolving patterns of surface temperature change in climate models *J. Clim.* **28** 1630–48

Armour K C, Bitz C M and Roe G H 2013 Time-varying climate sensitivity from regional feedbacks *J. Clim.* **26** 4518–34

Ceppi P and Gregory J M 2017 Relationship of tropospheric stability to climate sensitivity and Earth's observed radiation budget *Proc. Natl Acad. Sci.* **114** 13126–31

Dufresne J-L and Bony S 2008 An assessment of the primary sources of spread of global warming estimates from coupled atmosphere-ocean models *J. Clim.* **21** 5135–44

Eby M *et al* 2013 Historical and idealized climate model experiments: an intercomparison of Earth system models of intermediate complexity *Clim. Past* **9** 1111–40

Ehlert D and Zickfeld K 2017 What determines the warming commitment after cessation of CO_2 emissions? *Environ. Res. Lett.* **12** 015002

Frölicher T L and Paynter D J 2015 Extending the relationship between global warming and cumulative carbon emissions to multi-millennial timescales *Environ. Res. Lett.* **10** 075002

Frölicher T L, Winton M and Sarmiento J L 2014 Continued global warming after CO_2 emissions stoppage *Nat. Clim. Change* **4** 40–4

Geoffroy O, Saint-Martin D and Ribes A 2012 Quantifying the sources of spread in climate change experiments *Geophys. Res. Lett.* **39** L24703

Good P *et al* 2015 Nonlinear regional warming with increasing CO_2 concentrations *Nat. Clim. Change* **5** 138–42

Gregory J M, Andrews T and Good P 2015 The inconstancy of the transient climate response parameter under increasing CO_2 *Phil. Trans. R. Soc.* A **373** 20140417

Gregory J M, Ingram W J, Palmer M A, Jones G S, Stott P A, Thorpe R B, Lowe J A, Johns T C and Williams K D 2004 A new method for diagnosing radiative forcing and climate sensitivity *Geophys. Res. Lett.* **31** L03205

Hansen J, Lacis A, Rind D, Russell G, Stone P, Fung I, Ruedy R and Lerner J 1984 *Climate Sensitivity: Analysis of Feedback Mechanisms* (Washington, DC: American Geophysical Union) pp 130–63

IPCC 2013 *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* ed T F Stocker (New York: Cambridge University Press) p 1535

Jonko A K, Shell K M, Sanderson B M and Danabasoglu G 2013 Climate feedbacks in CCSM3 under changing CO_2 forcing: II. Variation of climate feedbacks and sensitivity with forcing *J. Clim.* **26** 2784–95

Kuhlbrodt T and Gregory J M 2012 Ocean heat uptake and its consequences for the magnitude of sea level rise and climate change *Geophys. Res. Lett.* **39** L18608

Matthews D H and Zickfeld K 2012 Climate response to zeroed emissions of greenhouse gases and aerosols *Nat. Clim. Change* **2** 338–41

Oliphant T E 2006 A Bayesian perspective on estimating mean, variance, and standard-deviation from data (https://scholarsarchive.byu.edu/facpub/278) (Accessed: 3 November 2018)

Paynter D and Frölicher T L 2015 Sensitivity of radiative forcing, ocean heat uptake, and climate feedback to changes in anthropogenic greenhouse gases and aerosols *J. Geophys. Res.: Atmos.* **120** 9837–54

Pfister P L 2017 Energy balance and ocean interactions in idealized future projections based on reduced-complexity climate models *PhD Thesis* (Climate and Environmental Physics) University of Bern 243

Pfister P L and Stocker T F 2017 State-dependence of the climate sensitivity in Earth system models of intermediate complexity *Geophys. Res. Lett.* **44** 10,643–10,653

Raper S C B, Gregory J M and Stouffer R J 2002 The role of climate sensitivity and ocean heat uptake on AOGCM transient temperature response *J. Clim.* **15** 124–30

Ritz S P, Stocker T F and Joos F 2011 A coupled dynamical ocean-energy balance atmosphere model for paleoclimate studies *J. Clim.* **24** 349–75

Rose B E J, Armour K C, Battisti D S, Feldl N and Koll D D B 2014 The dependence of transient climate sensitivity and radiative feedbacks on the spatial pattern of ocean heat uptake *Geophys. Res. Lett.* **41** 1071–8

Rose B E J and Rayborn L 2016 The effects of ocean heat uptake on transient climate sensitivity *Curr. Clim. Change Rep.* **2** 190–201

Roth R, Ritz S P and Joos F 2014 Burial-nutrient feedbacks amplify the sensitivity of atmospheric carbon dioxide to changes in organic matter remineralisation *Earth Syst. Dyn.* **5** 321–43

Siegenthaler U and Oeschger H 1984 Transient temperature changes due to increasing CO_2 using simple models *Ann. Glaciol.* **5** 153–9

Solomon S, Plattner G-K, Knutti R and Friedlingstein P 2009 Irreversible climate change due to carbon dioxide emissions *Proc. Natl Acad. Sci. USA* **106** 1704–9

Steinacher M and Joos F 2016 Transient Earth system responses to cumulative carbon dioxide emissions: linearities, uncertainties, and probabilities in an observation-constrained model ensemble *Biogeosciences* **13** 1071–103

Steinacher M, Joos F and Stocker T F 2013 Allowable carbon emissions lowered by multiple climate targets *Nature* **499** 197

Stocker B D, Roth R, Joos F, Spahni R, Steinacher M, Zaehle S, Bouwman L, Xu-Ri and Prentice I C 2013 Multiple greenhouse-gas feedbacks from the land biosphere under future climate change scenarios *Nat. Clim. Change* **3** 666–72

Stouffer R J 2004 Time scales of climate response *J. Clim.* **17** 209–17

Taylor K E, Stouffer R J and Meehl G A 2012 An overview of CMIP5 and the experiment design *Bull. Am. Meteorol. Soc.* **93** 485–98

Ullman D J and Schmittner A 2017 A cloud feedback emulator (CFE, version 1.0) for an intermediate complexity model *Geosci. Model Dev.* **10** 945–58

Winton M, Griffies S M, Samuels B L, Sarmiento J L and Froelicher T L 2013 Connecting changing ocean circulation with changing climate *J. Clim.* **26** 2268–78

Winton M, Takahashi K and Held I M 2010 Importance of ocean heat uptake efficacy to transient climate change *J. Clim.* **23** 2333–44